

PMC26

APPLICATION OF THE FRAMEWORK FOR EVALUATING COMPLEX INTERVENTIONS TO CLUSTER RANDOMIZED TRIALS FOR THE EVALUATION OF DISEASE MANAGEMENT PROGRAMS

Marchisio S¹, Panella M²

¹University Politecnica delle Marche, Ancona, Italy, ²Univerisity of Eastern Piedmont "A. Avogadro", Novara, NO, Italy

Trials of disease management programs pose several methodological challenges. Our objective is to assess the extent to which the various development steps of a cluster randomized trial to evaluate disease management are represented in the framework for the design and evaluation of complex interventions. The framework for evaluating complex interventions developed by Campbell and colleagues is composed by five phases: theoretical, identification of components of the intervention, definition of trial and intervention design, methodological issues for main trial, and promoting effective implementation. Using these phases the corresponding stages in the development of the cluster randomized trial to evaluate the effectiveness of disease management programs are identified and described. Synthesis of evidence needed to construct the program, survey and qualitative research used to define components of the program, a pilot study to assess the feasibility of delivering the care, methodological issues in the main trial including choice of design, allocation concealment, outcomes, sample size calculation and analysis are adequately represented using the stages of the framework for evaluating complex interventions. Even though is difficult to define precisely what exactly the "active ingredients" of a program of disease management and how they relate to each other, we think that the applied framework is a powerful resource for researchers planning a randomized clinical trial to evaluate the effectiveness of such programs.

WITHDRAWN

RESEARCH ON METHODS & CONCEPTUAL PAPERS—Patient-Reported Outcomes Studies

PMC28

RASCH PARTIAL CREDIT ANALYSIS OF THE SF-12V2 USING THE 2003 MEDICAL EXPENDITURE PANEL SURVEY (MEPS)

Gu NY, Doctor JN

University of Southern California, Los Angeles, CA, USA

This study assesses the Rasch measurement properties of the SF-12 version 2 (SF-12v2) physical and mental health (PH and MH) items in respondents with most prevalent chronic conditions. Medical Expenditures Panel Survey (MEPS) respondents' age ≥ 18 with complete SF-12v2s from 2003 were extracted ($n = 19,906$). Eleven subgroups were identified using the primary ICD-9-CM code for the top 10 chronic conditions (hypertension, diabetes, depression, back disorder, arthropathy, cholesterol, asthma, sinusitis, anxiety and joint disorder) and healthy persons ($n = 8324$). Respondents with perfect scores demonstrating ceiling ($n = 303$) and floor effects ($n = 12$) were removed to ensure uncertainty in the responses. Coding reflected that higher scores represent healthier respondents. The Rasch partial credit model was used to exam the item category properties and fit statistics. Residual factor analysis was conducted to assess the factor loadings on the items. Item misfit took place mostly among MH items (infit/outfit z -score > 2.0). Particularly, MH item "Have you felt calm and peaceful?" showed misfit across all subgroups. Rasch residual analysis identified 75.2% to 86.9% of the shared variance within clus-

ters of all items. Further, PH items had positive factor loadings while MH items had negative factor loadings, with few exceptions (3/11). Some item category steps had poor step properties. Mental health items are more likely to be noisy. The patterns of the factor loadings of the PH and MH items were as expected in most of the groups suggesting that the SF-12v2 has two distinct factors measuring the overall health. Poor item category performance suggests that collapsing of these categories might improve the quality of the instrument.

PMC29

INTERNATIONAL VALUATION SET FOR EQ-5D HEALTH STATES

Craig BM¹, Busschbach JJ², Salomon J³

¹Moffitt Cancer Center, Tampa, FL, USA, ²Erasmus Medical Center, Rotterdam, Netherlands, ³Harvard University, Boston, MA, USA

OBJECTIVE: Health states defined by classification systems like the EQ-5D can be valued using techniques that aim to elicit cardinal measures directly, such as the time trade-off (TTO) and visual analogue scale (VAS), but also by ranking the health states. In this study, we estimate international value sets for the 243 EQ-5D states based on rank, TTO and VAS responses and test their equivalence. **METHODS:** We estimated the coefficients scale is recommended in Sweden [in Swedish] of a conditional logit and a linear probability model of rank responses as well as the coefficients of two linear models of TTO and VAS responses using pooled data from eight countries: Slovenia, Argentina, Denmark, Japan, Netherlands, Spain, UK and US, which gave us 179,431 responses from 11,483 subjects. The main difference between the two rank models is that one models utility in $\ln(\text{odds})$, as suggested by McFadden, and the second in probability. The regression specifications have previous models estimated in the United States and UK nested within its framework. Furthermore, we compare rescaled predicted values for the 242 EQ-5D states, excluding 11,111, across techniques in terms of correlation and concordance. **RESULTS:** The non-optimal gap reduces when utilities are linearly modeled in probability instead of log-odds, as suggested by McFadden. The rank-based values are highly correlated with both TTO and VAS values. Compared to the log-odds model, the linear probability model produces rank-based values with greater concordance to TTO/VAS values. **CONCLUSION:** In former investigations we tested if ranking in the large pool of data produce values of health states comparable with direct measures as TTO and VAS. In this investigation we provide further evidence by showing convergent validity between TTO/VAS and alternative rank-based models for the whole valuation space. This evidence emphasizes the promise of a valuation technique that incorporates ordinal responses. Overall, rank-based values merits closer investigation.

PMC30

A REVIEW AND CRITIQUE OF METHODS FOR MEASURING TEMPORARY HEALTH STATES IN COST-UTILITY ANALYSES

Wright DR¹, Wittenberg E², Swan JS³, Miksad R⁴, Prosser L⁵

¹Harvard School of Public Health, Boston, MA, USA, ²Brandeis University, Waltham, MA, USA, ³Massachusetts General Hospital, Boston, MA, USA, ⁴Beth Israel Deaconess Medical Center, Boston, MA, USA, ⁵Harvard Medical School, Boston, MA, USA

OBJECTIVES: Temporary health states (states with a duration of less than one year) are common and include many infectious diseases, short-term treatments, and diagnostic procedures. Valuation of these states requires special consideration because the health state is transitory and standard methods ignore the influence of duration on preferences. Inaccurate assessments could introduce bias into cost-utility ratios. There is no "gold stan-